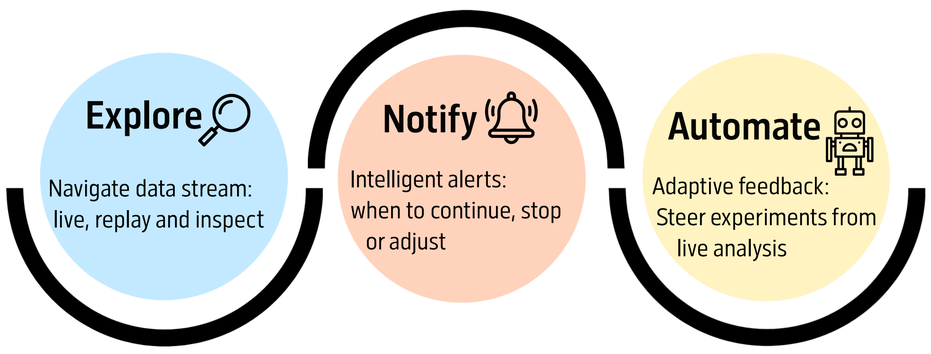
Intelligent Analysis Pipeline for X-ray Absorption Spectroscopy: Near Real-Time Processing and Adaptive Experimentation

Diana Rueda, Marc-Olivier Andrez, Tim Schoof, Mikhail Karnevskiy, Sergej Levcenco, Anuradha Bhogra, Martin Gasthuber, Edmund Welter, Anton Barty
Deutsches Elektronen-Synchrotron DESY, Hamburg, Germany

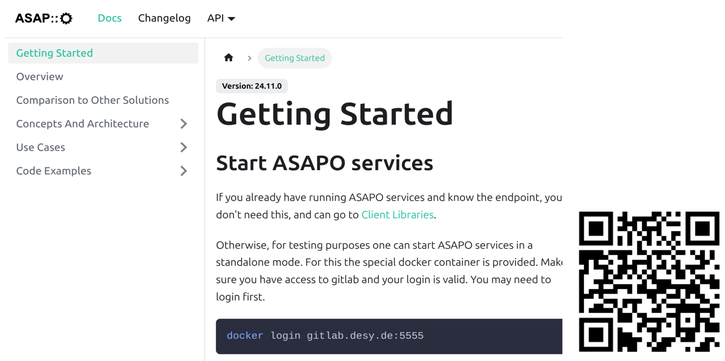
Enabling Responsive Experiments

Effective adaptive experimentation requires not only rapid analysis but also seamless integration of analytical results into the **experimental decision-making process**. This pipeline processes XAFS data with sub-second latency and provides multiple pathways for scientific insight.



Near Real-Time Processing

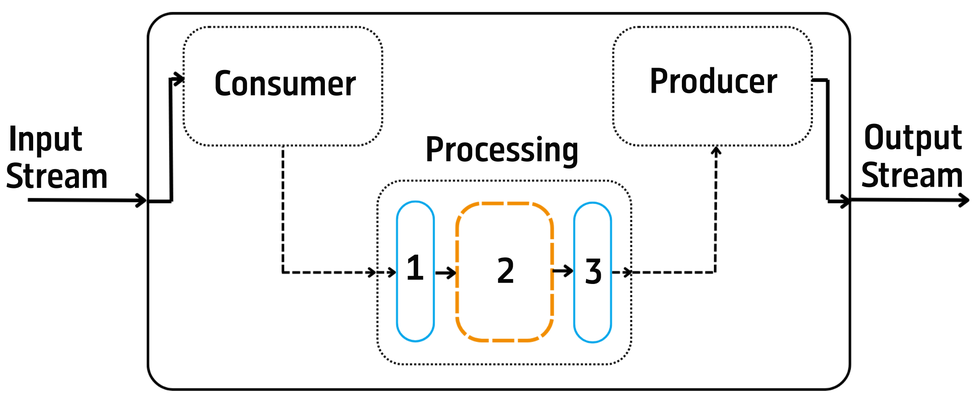
To address these challenges, DESY has developed **Asapo**, a **distributed streaming framework** designed to handle very large data payloads. Its main building block are the **Producers** that send data in message streams and **Consumers** that can subscribe to these streams an ingest the data. Asapo can be used with a standalone docker container and with any of its libraries in **Python**, **C++**, **C**, and bindings for **Haskell** and **Rust.**



Getting Started for the Asapo Streaming Network

The Analysis Worker

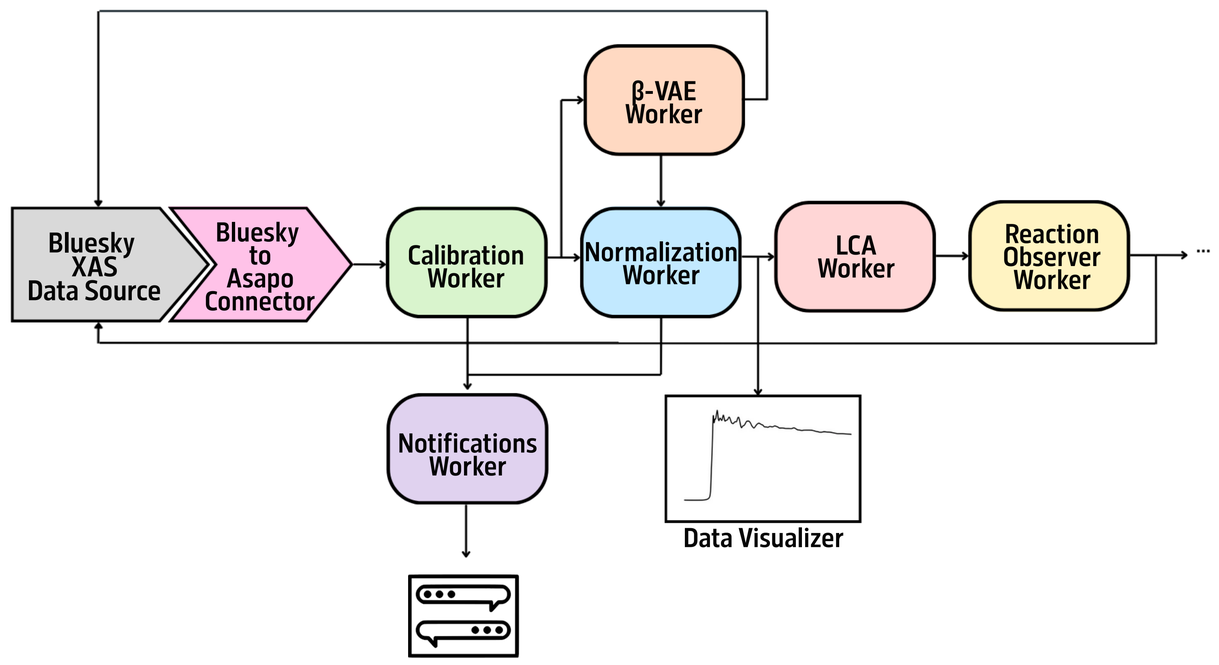
The **AsapoWorker** Python library is a higher-level abstraction that simplifies the construction of real-time data processing pipelines. It aims to **encapsulate diverse analytical processes** as standardized worker components that can be seamlessly integrated into processing pipelines.



Conceptual diagram for an Asapo Worker

The conceptual diagram above illustrates the **main logical components** of the AsapoWorker. Component 1 performs **deserialization** and 3 performs **serialization**. Component 2 represents the actual **analysis method in Python**, which can be adapted from conventional analysis approaches or ML approaches in inference mode.

Intelligent Analysis Pipeline



This is the analysis pipeline for the **beamline P65** for the **ROCK-IT project** at **DESY**. Its components can be interchangable, and integrate new approaches.

Data Source and Connector: The data sources can come directly from **fast detectors** or from a **control system** that collects the information from all detector. The connector translates this data into an **Asapo stream.**

Calibration Worker: It calibrates for the **energy shift** overtime with respect to a reference sample. It can be extended to **predict this energy drift** through **ML approaches** for the cases when there is not enough signal to measure a reference sample.

β-VAE Worker: Based on analysis developed by **Gregor Hartmann from Helmholtz-Zentrum Berlin**. It reconstructs spectra and extracts disentangled features that could inform on meaningful changes in the experiment.

Normalization Worker: It uses the **Larch library** to perform pre-edge fitting, post-edge fitting, edge step calculation and normalization.

LCA Worker: It is also used the the Larch library. As output the fitted absorption coefficient as well as the **weights for the contribution** of each of the references in this fit.

Reaction Observer Worker: This worker can suggest **when to stop measurements in an experiment stage** based on how stable is the oxidation level in a predetermined amount of time.

Notifications Worker: Collects notifications sent by other workers in the pipeline and it **publishes them in a chat**.

Data Visualizer: Web-based visualization tool that generically plots data stored in **Asapo** in **near real-time**.

Would you like to include your analysis in this pipeline? Let's discuss!

AsapoWorker https://gitlab.desy.de/fs-sc/asapoworker
Asapo https://asapo.pages.desy.de/asapo/
Xdi Pipeline https://gitlab.desy.de/fs-sc/analysis-pipelines/xdi_pipeline
Asapo Bluesky Connector https://gitlab.desy.de/fs-sc/analysis-pipelines/asapo_bluesky_connector
Questions/comments diana.rueda@desy.de, marc-olivier.andrez@desy.de, tim.schoof@desy.de,
We gratefully acknowledge funding from the Helmholtz Association HGF for the ROCK-IT project.

ROCK-IT

ASAP::

HELMHOLTZ